TECHNIQUES FOR SCRATCH AND DATE REMOVAL FROM SCANNED FILM

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to techniques for removing from scanned film dates and/or missing data regions that can be the result of spots, scratches or folds on the film. The techniques of the present invention, which may be implemented in an apparatus and/or as methods, are particularly designed to handle larger regions of missing data, such as "thick" scratches. The invention also relates to programs of instructions for directing an apparatus or machine to carry out these techniques.

2. Description of the Related Art

Current methods for removing scratches from scanned film are designed for thin scratches which are only a few pixels wide. The removal of these artifacts typically involve either median or mean filtering. However, there are drawbacks with both of these methods when larger missing data regions are involved. The problem is that a median filter is sensitive only to the color histograms of pixels in a neighborhood, not to the spatial distribution of these pixels. Therefore, application of a median filter on a larger region of missing data tends to produce "blocky" artifacts. Mean filtering, on the other hand, tends to blur regions and boundaries in the process of attempting to remove larger missing data regions.

OBJECTS AND SUMMARY OF THE INVENTION

Objects of the Invention

Therefore, it is an object of the present invention to overcome the aforementioned problems.

It is another object of this invention to provide techniques for identifying missing data regions (which may include characters in a date field) and performing multiple filtering operations, one of which is performed using a "closest to radial basis function" (CRBF) approach.

SUMMARY OF THE INVENTION

One aspect of this invention involves a method for processing data in at least one portion of film that has been scanned to generate a pixel representation. The method comprises the steps of: segmenting at least one portion of the pixel representation to identify at least one region of missing data; calculating an area/perimeter ratio for each identified region of missing data; and subjecting each identified region of missing data having an area/perimeter ratio less than a predetermined maximum to a closest-to-radial-based-function filtering operation to estimate pixel values in that region from neighboring pixel values.

Preferable aspects of the segmenting include mapping the three dimensional color space of the pixels in each portion to a one dimensional line segment. The segmenting also preferably includes establishing a reference color for each portion; and determining a corresponding anchor color based on the established reference color; wherein the one dimensional line segment is defined by the reference color at one end and the anchor color at the other end. The segmenting still further preferably includes quantizing the one dimensional line segment into a plurality of bins, each of which is identified with a bin-index; creating a co-occurrence matrix M[i][j] for each portion, M[i][j] being equal to the number of pixel locations in that portion, such that a current pixel has bin-index i and its right or bottom neighbor has bin-index j; selecting a threshold that creates two areas in the co-occurrence matrix and that maximizes the entropy of the data in each of the two areas; and identifying pixels having a bin-index greater than the threshold as missing data.

Preferably, the calculating comprises performing a component filtering operation.

Preferably, for each pixel value estimated, the closest-to-radial-based-function filtering operation comprises using color values of neighboring pixels, without introducing any new colors, to estimate that pixel value and to fill each region of missing data. The closest-to-radial-based-function filtering operation considers spatial distribution and color distribution information in estimating pixel values to fill each region of missing data.

According to another aspect of the invention, an apparatus is provided for processing data in at least one portion of a pixel representation of film that has been scanned to generate a pixel representation. The apparatus comprises: a segmentation module that segments at least a portion of the pixel representation to identify at least one region of missing data; a calculator that calculates an area/perimeter ratio for each identified region of missing data; and a closest-to-radial-based-function filter that subjects each identified region of missing data having an area/perimeter ratio less than a predetermined maximum to a closest-to-radial-based-function filtering operation to estimate pixel values in that region from neighboring pixel values.

Preferable features of the segmentation module include mapping the three dimensional color space of the pixels in each portion to a one dimensional line segment. The segmentation module also preferably establishes a reference color for each portion; and determines a corresponding anchor color based on the established reference color; wherein the one dimensional line segment is defined by the reference color at one end and the anchor color at the other end. The segmentation module still further preferably quantizes the one dimensional line segment into a plurality of bins, each of which is identified with a bin-index; creates a co-occurrence matrix M[i][j] for each portion, M[i][j] being equal to the number of pixel locations in that portion, such that a current pixel has bin-index i and its right or bottom neighbor has bin-index j; selects a threshold that creates two areas in the co-occurrence matrix and that maximizes the entropy of the data in each of the two areas; and identifies pixels having a bin-index greater than the threshold as missing data.

Preferably, the calculator performs a component filtering operation.

Preferably, for each pixel value estimated, the closest-to-radial-based-function filtering operation comprises using color values of neighboring pixels, without introducing any new colors, to estimate that pixel value and to fill each region of missing data. The closest-to-radial-based-function filtering considers spatial distribution and color distribution information in estimating pixel values to fill each region of missing data.

In accordance with further aspects of the invention, any of the above-described methods or steps thereof may be embodied in a program of instructions (e.g., software) which may be stored on, or conveyed to, a computer or other processor-controlled device for execution. Alternatively, any of the methods or steps thereof may be implemented using